A SCALABLE APPROACH TO MODELING CASCADING RISK IN THE MDAP NETWORK

Anita Raja, Mohammad Rashedul Hasan, Shalini Rajanna
Department of Software and Information Systems
University of North Carolina at Charlotte

Ansaf Salleb-Aoussi
Center for Computational Learning Systems
Columbia University

Acknowledgements: Dr. Maureen Brown UNCC, Mr. Rob Flowe AT&L, Jagan Vujjini, UNCC
This material is based upon work supported by the Naval Postgraduate School Acquisition Research Program under Grant No. N00244-11-1-0024

UNC CHARLOTTE
STAKE YOUR CLAIM
WWW.UNCC.EDU

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
MAY 2014		**00-00-2014 to 00-00-2014**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A Scalable Approach to Modeling Cascading Risk in the MDAP Network	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of North Carolina at Charlotte,Department of Software and Information Systems,Charlotte,NC,28223**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
AFCEA 11th Annual Acquisition Research Symposium, 14-15 May 2014, Monterey, CA.

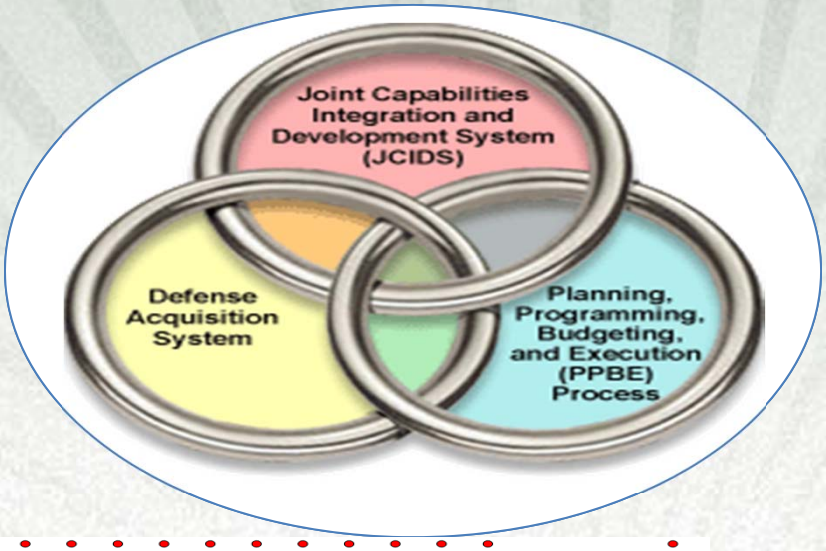
14. ABSTRACT

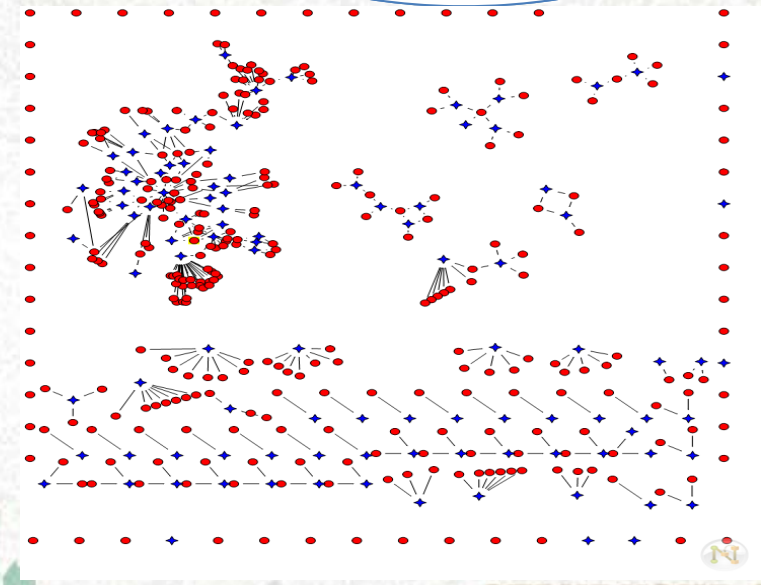
15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**23**	

MOTIVATION

- Joint Capabilities:
  - Future operating environment: uncertainty, complexity, rapid change, and persistent conflict.
  - Requires integrated approach (WSARA 2009)
- Growing PE / MDAP interdependencies and complexity.
  - Focus on funding interdependency



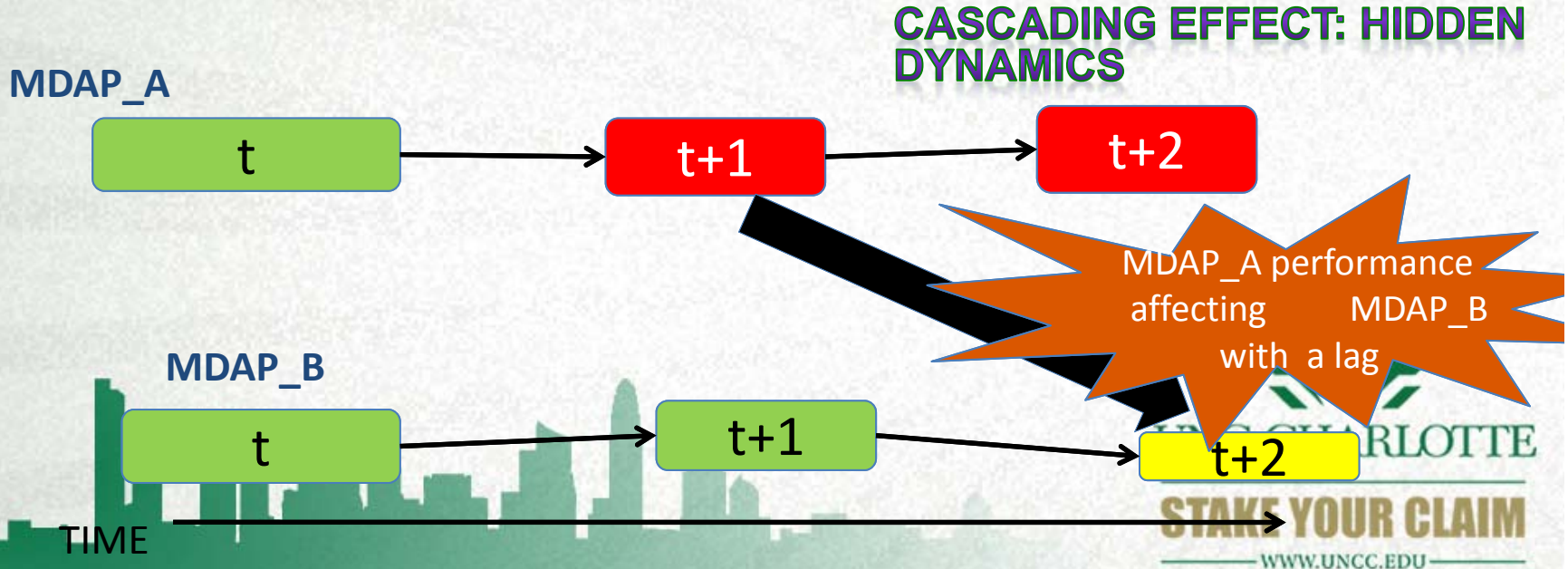


RESEARCH PATHS

- Network-centric approach (Brown and Owen, 2012; Raja et al., 2012)
- Automated analysis for correlation and associations (Zhao et al, 2012)
- Portfolio-based approach (Davendralingam et al.)

OUR PREVIOUS WORK (RAJA ET AL.,2012, 2013):

- Data-driven approach to develop what-if models to predict early indicators of cascading risks.
- DAES reports of small set of MDAPs over several years.
- Results:
  - Non-local factors affect program outcomes: “program-centric” + “program network approach” for acquisition and management.
  - Cascading effects recast as a sequential decision problem.
  - Tedious manual process to build decision process model.

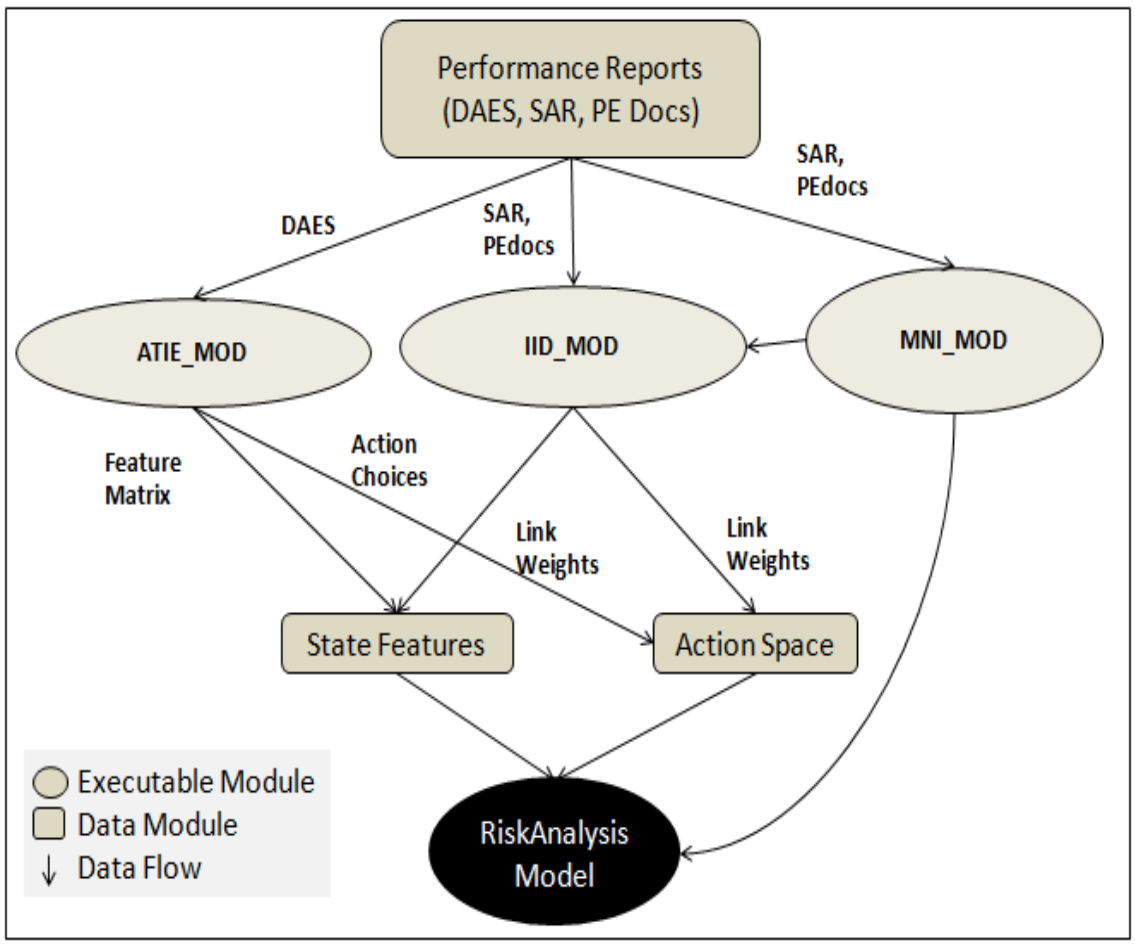


CURRENT RESEARCH GOAL

- Automate Extraction & Analysis of MDAP DAES PSM, Issues, Actions:
  - DAES data for multiple programs over multiple years.
- Automate Extraction of Structural Properties of MDAP Network:
  - Identify funding neighbors based on PE and SAR data.
  - Determine link weights from PE perspective.
- Populate Decision Process Model.
- Identify challenges to data acquisition.

UNC CHARLOTTE
STAKE YOUR CLAIM
WWW.UNCC.EDU

DECISION SUPPORT FRAMEWORK



Legend:

ATIE_MOD Automated Text & Image Extraction Module

IID_MOD Interdependency Determiner Module

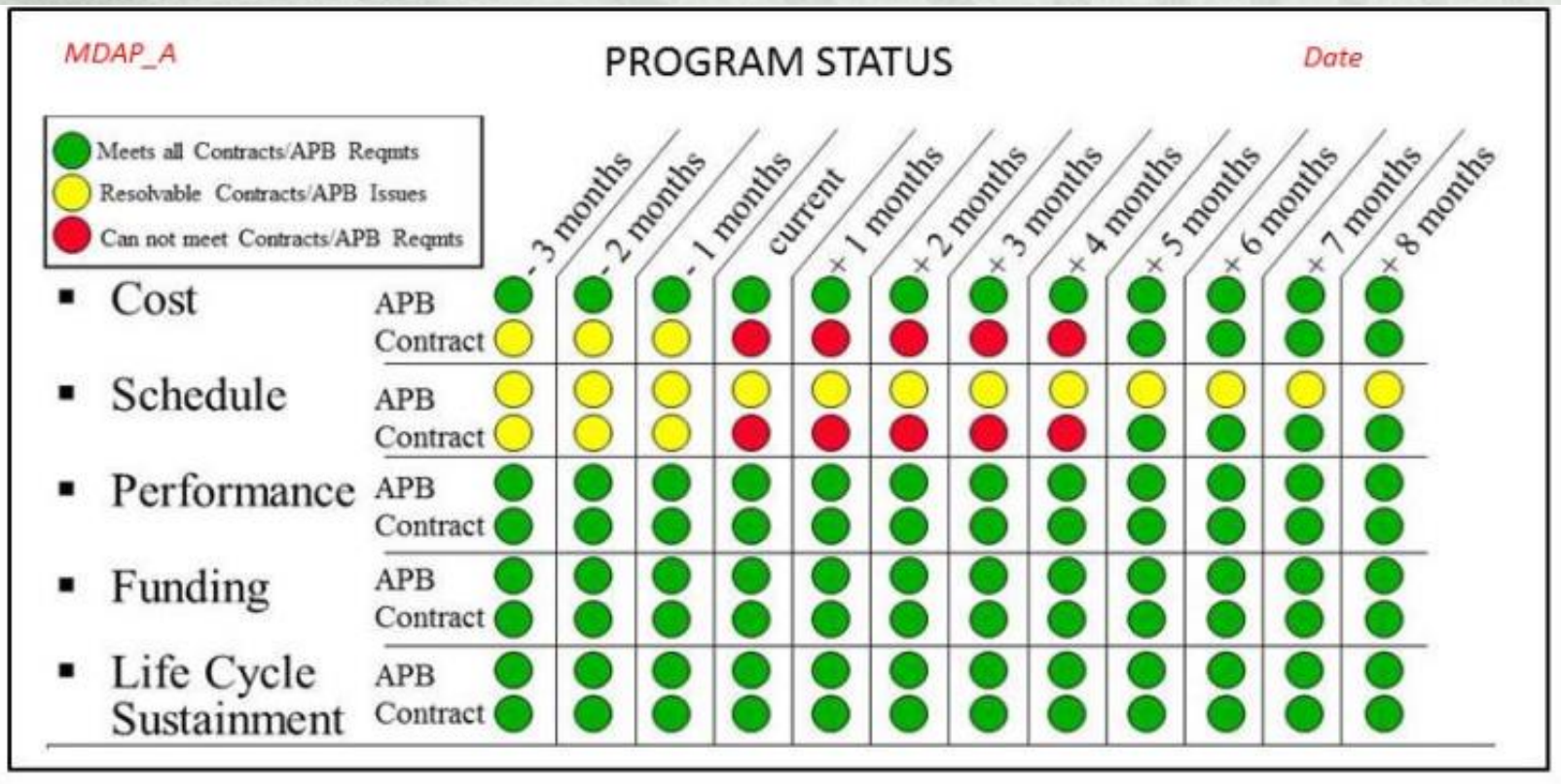
MNI_MOD MDAP Network Identifier Module

UNC CHARLOTTE
STAKE YOUR CLAIM
WWW.UNCC.EDU

MAIN CONTRIBUTIONS

- Network and Program Centric analysis.
- Novel Integration of methodologies for text and image analytics for large scale automated data extraction.
- Quantification of interdependency metric.
- Progress towards Decision support framework for MDAP cascading risk prediction.

INPUT DATA : PROGRAM STATUS MATRIX



Green- Contracts that are met,
Yellow- Resolvable Contracts
Red- Contracts that cannot be resolved

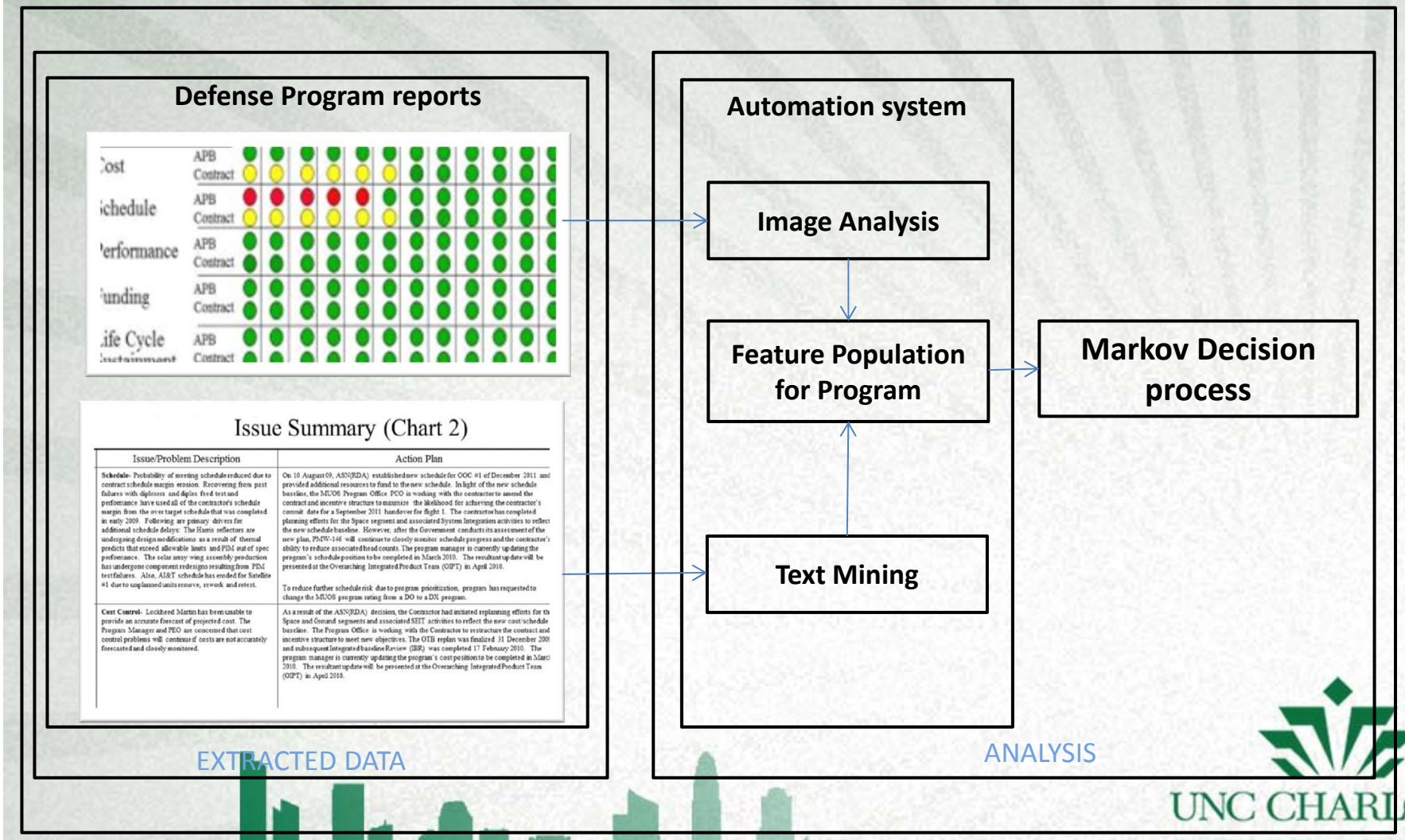
INPUT DATA: ISSUE SUMMARY – 4 SAMPLE RECORDS

ISSUE SUMMARY	
ISSUES	ACTIONS
Schedule-The reflectors are undergoing design modifications as a result of thermal predicts that exceed allowable limits and PIM out of spec performance	Program Office is working to amend the contract and the incentive structure to maximize the likelihood for achieving the contractor's commit date
Cost Control- A Team, has been unable to provide an accurate forecast of projected cost.	XYZ has detailed manpower projections and is measuring each organization's effectiveness in reducing headcount, a metric considered key to managing costs

MDP STATE FEATURE MATRIX

Feature Number	Feature Description	An example of Feature value
Feature 1	Program ID	MDAP_A
Feature 2	Current Year	2010
Feature 3	Current Month	April
Feature 4	Cost(APB) status, for 9 months starting with current month	<1,1,1,1,1,1,1,1,1,1,1,1>
Feature 5	Cost(Contract) status, for 9 months starting with current month	<0,0,0,-1,-1,-1,-1,-1,1,1,1,1>
Feature 6	Schedule(APB) status, for 9 months starting with current month	<0,0,0,0,0,0,0,0,0,0,0,0>
Feature 7	Schedule(Contract)status, for 9 months starting with current month	<0,0,0,-1,-1,-1,-1,-1,1,1,1,1>
Feature 8	Performance (APB) status, for 9 months starting with current month	<1,1,1,1,1,1,1,1,1,1,1,1>
Feature 9	Performance(Contract) status, for 9 months starting with current month	<1,1,1,1,1,1,1,1,1,1,1,1>
Feature 10	Funding(APB) status, for 9 months starting with current month	<1,1,1,1,1,1,1,1,1,1,1,1>
Feature 11	Funding(Contract) status, for 9 months starting with current month	<1,1,1,1,1,1,1,1,1,1,1,1>
Feature 12	List of Issues	Cost unable to forecast
Feature 13	List of Actions	Reduce resources

AUTOMATION OF TEXT/IMAGE ANALYSIS



LATENT DIRICHLET ALLOCATION(LDA)

- A Document is a bag of words; random mixture of latent topics
- Unsupervised learning - uncover the latent topics characterized by a statistical distribution in a given set of document.
- Result in distribution of 1) topics across documents and 2) words across topics.
- Training data: DAES reports of MDAP_A from 2007-2011 with 150 records

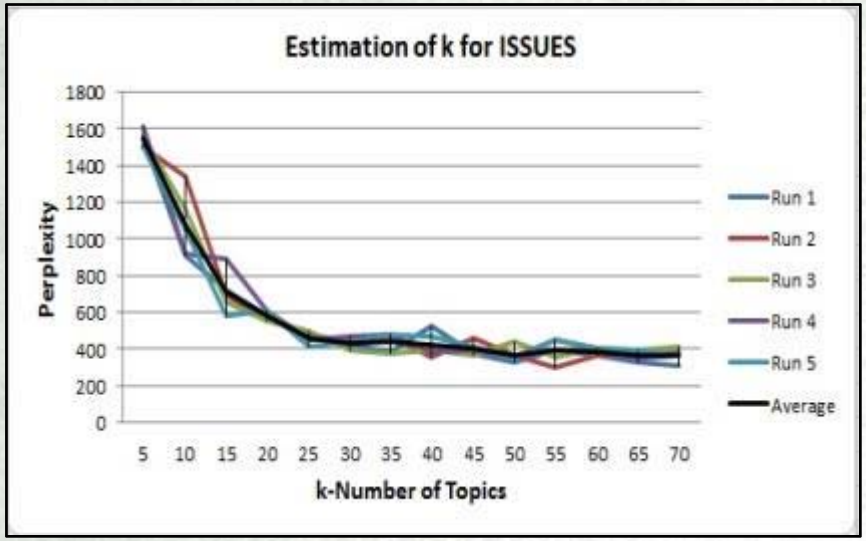
Farmers who received some government information late have been given additional time to file their 1987 federal income tax return. The internal revenue service said Monday it was granting special relief to farmer's who did not receive agriculture department documents by the Feb 15 scheduled payments and Form 1099-A, on which acquisitions or abandonments of secured property are reported.

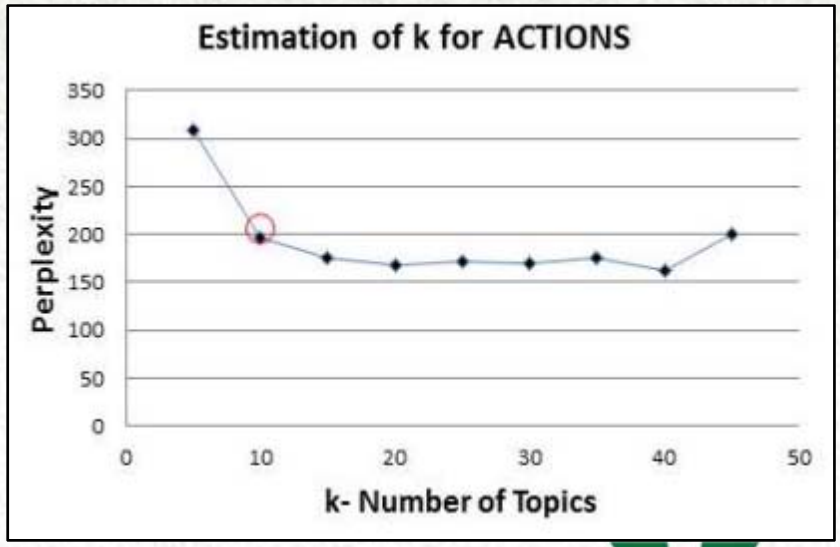
Topic 0	Topic 1	Topic 2	Topic 3
Tax	last	cents	office
Money	week	future	department
Income	month	lower	general
Paid	days	farmers	investigation
Pay	april	cent	attorney
Trust	weeks	higher	justice
Fund	came	tons	government
Taxes	june	drought	law

**David M. Blei, Andrew Y. Ng and Michael I. Jordan, Latent Dirichlet Allocation, *Published in the Journal of Machine Learning Research*, Volume 3, Pages 993-1022, March 2003

MODEL PARAMETER ESTIMATION

- Perplexity is a measure of model's ability to infer the topics in unseen documents.
  - K= 15 for issues; 10 for actions.



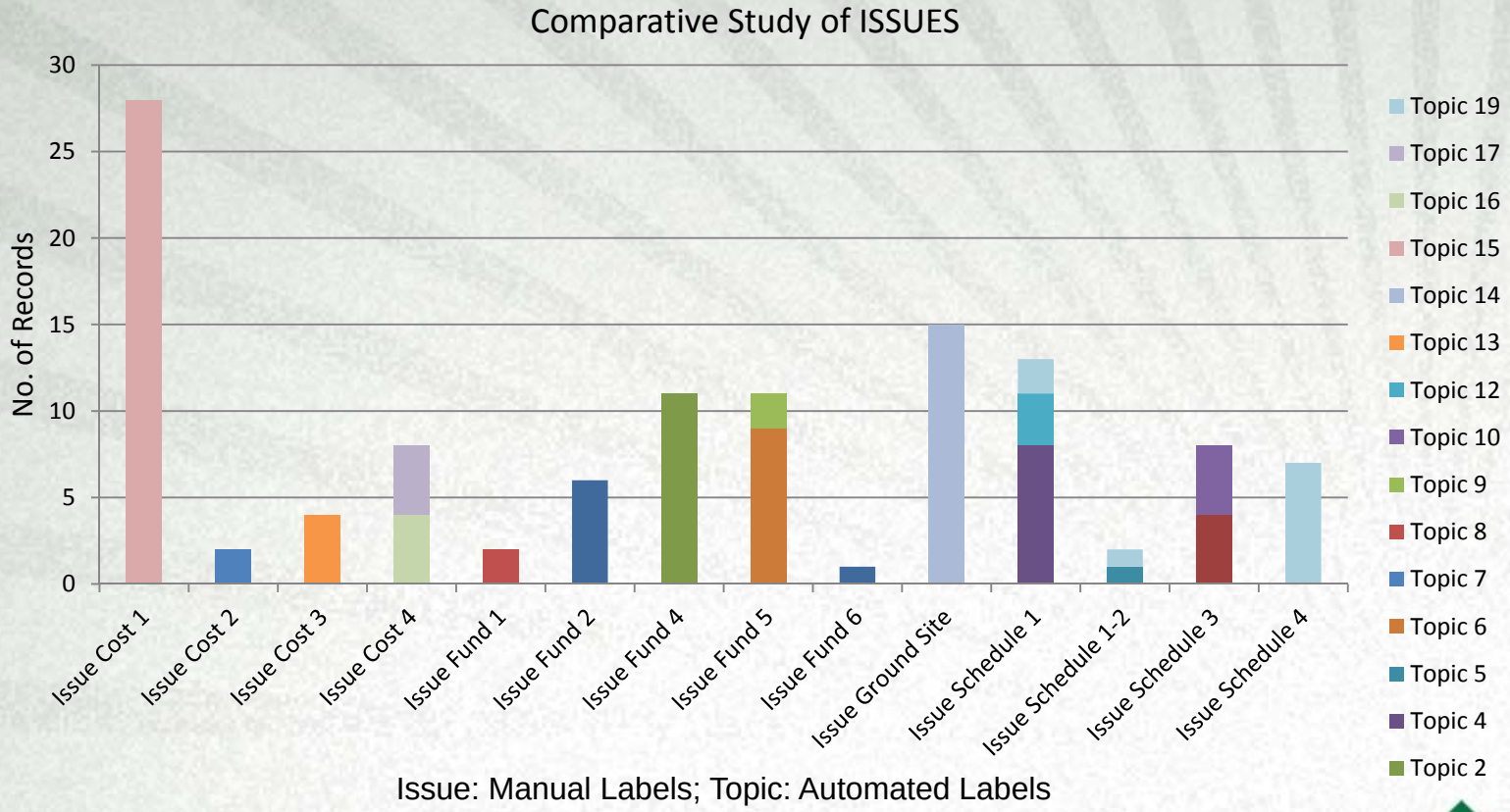


ISSUE AND ACTION TYPES

ISSUE	Issue Type	Description
Cost	Type 1	Repeated inability to accurately forecast cost
	Type 2	Overrun costs due to technical delays
	Type 3	Modifications to total cost to avoid going over budget
	Type 4	Overrun costs due to unanticipated extra elements in the project
Schedule	Type 1	Hardware Related Issues
	Type 2	Software Related Issues
	Type 3	International Collaboration Delay
	Type 4	Execution Delay due to failure to meet standards
	Type 5	Funding Delay
Logistics	Type 1	Ground Site is unavailable for installation.
Fund	Type 1	Funding has been aligned with estimate
	Type 2	Funding is short
	Type 3	Funding has been frozen
	Type 4	Funding is being reassessed due to errors or changes in need

ACTION	Issue Type	Description
Cost	Type 1a	Contractor Renegotiation
	Type 1b	Engage with contractor to address cost issues
	Type 2	Monitor performance
	Type 3	Manage cost by reducing head count
Schedule	Type 1a	Contractor examines schedule for the program to additional margin
	Type 1b	Amend contract
	Type 2a	Projection of schedule based on hardware and I&T progress
	Type 2b	Update schedule based on hardware and I&T progress
	Type 3	APB baseline change
	Type 4	Investigate alternate sites
	Type 5	Accelerate schedule for early contract reward
	Type 7	Accelerate schedule for Early contract reward
Funding	Type 1	Keep Navy and OSD informed of progress
	Type 2	Work with OSD and senior Leadership
	Type 3	Fund DoD Teleports
	Type 4	Monitor program to determine when to launch fund
	Type 5	Work with Air Force Space Command (AFSPC)

AUTOMATED ANALYSIS : ISSUES



Issue Cost 1 : Repeated inability to accurately forecast cost : Topic 15

Issue Schedule 1: Hardware related issue: Topic 10, 12 19

UNC CHARLOTTE
STAKE YOUR CLAIM
WWW.UNCC.EDU

AUTOMATED ANALYSIS : ACTIONS

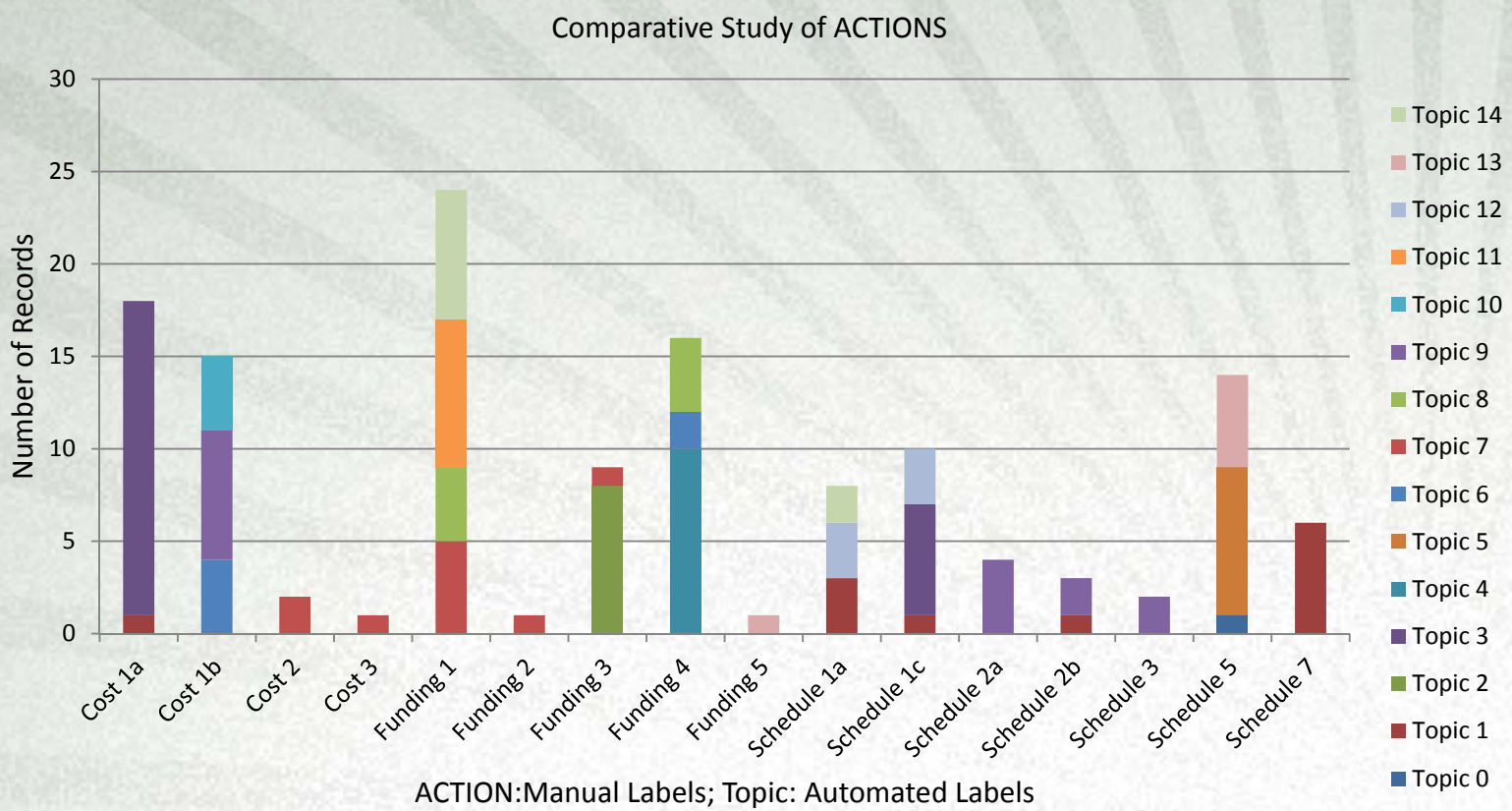
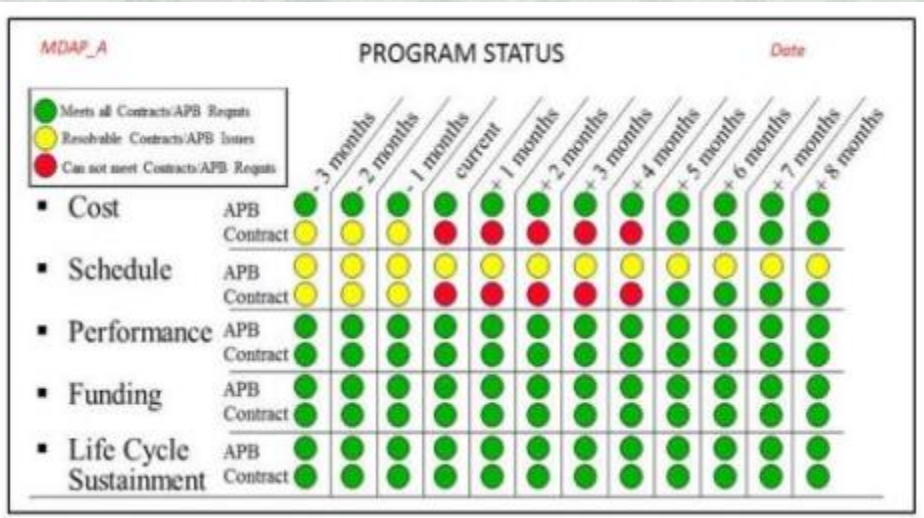


IMAGE ANALYSIS

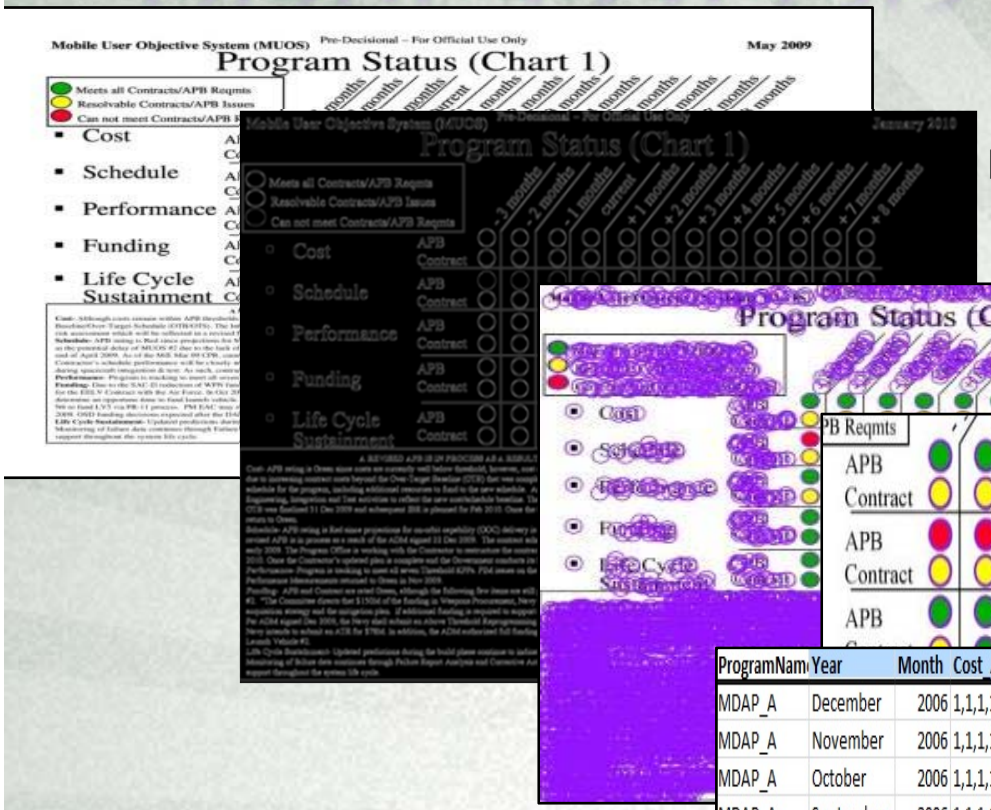
Conversion of program Status matrix to numeric format,
Green -> 1
Yellow -> 0
Red -> -1



Item	P3	P2	P1	Current	F1	F2	F3	F4	F5	F6	F7	F8
Cost_APB	1	1	1	1	1	1	1	1	1	1	1	1
Cost_Contract	0	0	0	-1	-1	-1	-1	-1	1	1	1	1
Schedule_APB	0	0	0	0	0	0	0	0	0	0	0	0
Schedule_Contract	0	0	0	-1	-1	-1	-1	-1	1	1	1	1
Funding_APB	1	1	1	1	1	1	1	1	1	1	1	1
Funding_Contract	1	1	1	1	1	1	1	1	1	1	1	1
Life cycle Sustainment_APB	1	1	1	1	1	1	1	1	1	1	1	1
Life cycle Sustainment_Contract	1	1	1	1	1	1	1	1	1	1	1	1

OpenCV –Open source Computer Vision software, java libraries is used for image analysis in this work.

IMAGE ANALYSIS ALGORITHM



Pre-processing of Image

Contour of Image

Process Circles

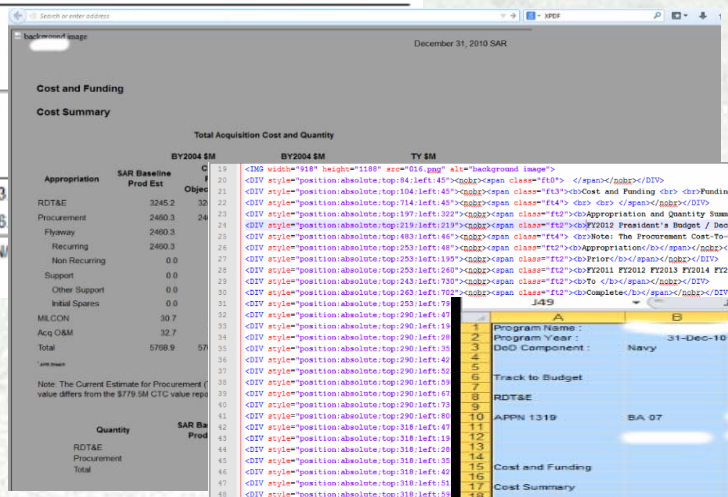
ProgramNam	Year	Month	Cost_APB	Cost_Cont	Schedule_APB	Schedule_Cont	Performance_APB	Performance_Cont
MDAP_A	December	2006	1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,	
MDAP_A	November	2006	1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,	
MDAP_A	October	2006	1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,	
MDAP_A	September	2006	1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,	
MDAP_A	April	2007	1,1,1,1,1,1,1,1,1,1,1,	0,0,0,0,0,0,0,0,0,0,0,	1,0,0,0,1,1,1,1,1,1,1,	1,0,0,0,1,1,1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,	0,0,0,0,0,0,0,0,0,0,0,
MDAP_A	February	2007	1,1,1,1,1,1,1,1,1,1,1,		1,1,1,0,0,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,	
MDAP_A	January	2007	1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,		1,1,1,1,1,1,1,1,1,1,1,	
MDAP_A	June	2007	1,1,1,1,1,1,1,1,1,1,1,	0,0,0,0,0,0,0,0,0,0,0,	0,0,1,1,1,1,1,1,1,1,1,	0,0,0,0,0,0,1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,	0,0,0,0,0,0,0,0,0,0,0,
MDAP_A	November	2007	0,0,0,0,1,1,1,1,1,1,1,	-1,-1,-1,-1,0,0,0,0,0,0,0,	1,1,0,0,1,1,1,1,1,1,1,	0,0,0,0,1,1,1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,	0,0,0,0,1,1,1,1,1,1,1,
MDAP_A	October	2007	1,0,0,0,0,1,1,1,1,1,1,	0,-1,-1,-1,-1,0,0,0,0,0,0,	1,1,1,0,0,1,1,1,1,1,1,	0,0,0,0,1,1,1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,	0,0,0,0,0,1,1,1,1,1,1,
MDAP_A	September	2007	1,1,0,0,0,0,0,0,0,0,0,	0,0,-1,-1,-1,-1,-1,-1,-1,-1,-1,	1,1,1,1,1,1,1,1,1,1,1,	0,0,0,0,0,1,1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,	0,0,0,0,0,0,0,0,0,0,0,
MDAP_A	April	2008	0,0,1,1,1,1,1,1,1,1,1,	0,0,0,0,0,0,0,0,0,0,0,	0,0,0,0,0,0,0,0,0,0,0,	0,0,0,0,0,0,0,0,0,0,0,	1,1,1,1,1,1,1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,
MDAP_A	August	2008	1,1,1,1,1,1,1,1,1,1,1,	0,0,-1,-1,-1,-1,-1,1,1,1,1	0,0,0,0,0,0,0,0,0,0,0,	0,0,-1,-1,-1,-1,-1,1,1,1,1	1,1,1,1,1,1,1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,
MDAP_A	December	2008	1,1,1,1,1,1,1,1,1,1,1,	-1,-1,-1,-1,-1,-1,1,1,1,1,	0,-1,-1,-1,-1,-1,1,1,1,1,1	-1,-1,-1,-1,-1,-1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,
MDAP_A	July	2008	1,1,1,1,1,1,1,1,1,1,1,	0,0,0,-1,-1,-1,-1,1,1,1,1	0,0,0,0,0,0,0,0,0,0,0,	0,0,0,-1,-1,-1,-1,1,1,1,1	1,1,1,1,1,1,1,1,1,1,1,	1,1,1,1,1,1,1,1,1,1,1,

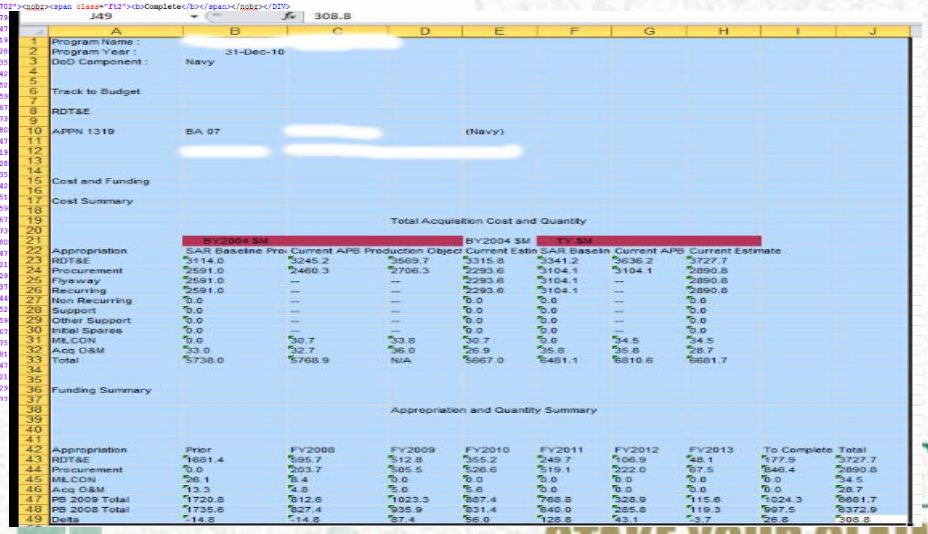
SAR EXTRACTION: COST SUMMARY, FUNDING SUMMARY, TRACK TO BUDGET

Total Acquisition Cost and Quantity

Appropriation	BY2004 $M SAR Baseline Prod Est	BY2004 $M Current APB Production Objective/Threshold	BY2004 $M Current Estimate	TY $M SAR Baseline Prod Est	TY $M Current APB Production Objective	TY $M Current Estimate
RDT&E	3245.2	3245.2	3569.7	3601.2	3636.2	3636.2	4040.0
Procurement	2460.3	2460.3	2706.3	2300.1	3104.1	3104.1	2831.3
Flyaway	2460.3	–					
Recurring	2460.3	–					
Non Recurring	0.0	–					
Support	0.0	–					
Other Support	0.0	–					
Initial Spares	0.0	–					
MILCON	30.7	30.7	33				
Acq O&M	32.7	32.7	36				
Total	5768.9	5768.9	N/				

APB Breach





PE Extraction

RDT&E appropriations

COST ($ in Millions)	All Prior Years	FY 2012	FY 2013#	FY 2014 Base	FY 2014 OCO ##	FY 2014 Total	FY 2015	FY 2016	FY 2017	FY 2018	Cost To Complete	Total Cost
Total Program Element	4,174.496	258.811	188.482	66.231	-	66.231	33.188	24.324	7.434	22.180	Continuing	Continuing
0728: EHF SATCOM Terminals	586.077	17.476	31.731	21.077	-	21.077	19.502	13.693	0.000	14.557	Continuing	Continuing
0731: FLTSATCOM	15.209	4.155	10.828	9.202	-	9.202	5.210	3.469	0.000	0.000	0.000	48.073
2472: Mobile User Objective Sys (MUOS)	3,573.210	237.180	145.5									

COST ($ in Millions)	All Prior Years	FY 2012	FY 2013#	FY 2014 Base	FY 2014 OCO ##	FY 2014 Total	FY 2015	FY 2016	FY 2017	FY 2018	Cost To Total	Complete Cost
Total Program Element	4,174.496	258.811	188.482	66.231	-	66.231	33.188	24.324	7.434	22.180	Continuing	Continuing
0728: EHF SATCOM Terminals	586.077	17.476	31.731	21.077	-	21.077	19.502	13.693	0	14.557	Continuing	Continuing
0731: FLTSATCOM	15.209	4.155	10.828	9.202	-	9.202	5.210	3.469	0	0.000	0	48.073
2472: Mobile User Objective Sys (MUOS)	3,573.210	237.180	145.923	35.952	-	35.952	8.476	7.162	7.434	7.623	130.912	4153.872

Excel Format

LESSONS LEARNED

- Text and image extraction/analysis:
  - Existing tools and technologies can support large scale automated analysis of DAES, PE docs, SARS
  - Topic models built from MDAP hub data seem to be relevant to neighbors.
  - Challenges: Formatting and Content inconsistencies; Missing data reports.
- Funding dependency measure can be obtained from PE and SAR data:
  - Funding amounts for the MDAPs as captured from the PE documents and respective SARs sometimes don't match
  - Funded MDAPs are not always listed in the Funding Summary page of PE: e.g., can be siphoned through a non-MDAP.

FUTURE WORK

- Large-scale extraction of entire set of available MDAP data.
- Populate Decision-theoretic model
  - State features, Action Space, Probability Transitions and Reward Functions (G,Y,R)
- Run “what-if” simulations.
- Extend to Data Network and MAIS networks.


UNC CHARLOTTE
STAKE YOUR CLAIM
WWW.UNCC.EDU